УДК 004.93

О.С. Коваль
Інститут кібернетики імені В.М. Глушкова Національної академії наук України, Україна
пр. Академіка Глушкова, 40, м. Київ, 03680
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського», Україна
пр. Перемоги, 37, корпус 18, м. Київ, 03056

ПРОМІЖНА СТАДІЯ ПОДІБНОГО ДО ЛЮДСЬКОГО ШТУЧНОГО ІНТЕЛЕКТУ

O.S. Koval
V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine, Ukraine
40, Academician Glushkov Ave., Kyiv, 03680
National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Ukraine
37, Victory Ave., Building, 18, Kyiv, 03056

AN INTERMEDIATE STAGE OF ARTIFICIAL INTELLIGENCE SIMILAR TO HUMAN

Проведено аналіз шляху подальшої розробки штучного інтелекту, подібного до людського, а саме його проміжної стадії, стану, коли він зможе вільно спілкуватись з людьми, зможе реалістично імітувати деяку задану особистість, вільно змінювати її і, тим самим, стане зручним для широкомасштабного використання у побуті. Розписано приблизний вигляд системи, яким він може стати у найближчому майбутньому з огляду на технології, що з'явились за останній час. Наведено перелік можливих налаштувань, що будуть актуальні для подібних систем. Перераховано можливі сфери застосування даного штучного інтелекту.

Ключові слова: штучний інтелект, напрямок розвитку, налаштування, сфери використання, імітація особистості, швидка заміна особистостей, штучний співрозмовник, чат-боти

An analysis of the way of further development of artificial intelligence similar to human was carried out, namely its intermediate stage, a state when he can communicate freely with people, can realistically imitate a some personality, freely change it, and thus become convenient for large-scale use in everyday life. An approximate view of the system was described, which it may become in the near future, given the recent technologies. The list of possible settings, which will be actual for similar systems, is resulted. The possible areas of application of this artificial intelligence are listed.

Keywords: artificial intelligence, direction of development, settings, areas of use, imitation of personality, quick replacement of personalities, artificial interlocutor, chatbots

Вступ

Обчислювальні можливості сучасних комп'ютерів з кожним роком невблаганно ростуть, а з ними збільшуються й можливості штучного інтелекту. Ще не відлунали перші гучні заяви про проходження тесту Тюрінга [1], а розробники демонструють людству все більш і більш досконалі чат-боти [2], вміння вести осмислену розмову яких викликає все менше запитань. Штучний інтелект потроху навчають не тільки імітувати відповіді людини, а й реально розмірковувати, вирішувати логічні задачі, що постають під час спілкування, відігравати деяку роль. На фоні значних досягнень у розробці нейронних мереж (наприклад, з нових публікацій можна виділити технології проведення заміни облич у потоці відео [3], клонування голосу [4], так звані діпфейки, видалення тіней з фотографій [5], прибирання перешкод із зображень [6], створення «штучних фотографій» з простих малюнків [7], створення фотографій людей, які насправді не існують [8], цікаві роботи в галузі математичного моделювання живої нейромережі людини, що проводяться у нашій країні [9,10] тощо)

стає очевидно, що ми, можливо, знаходимося ще далеко до моменту створення повноцінного, подібного до людського, штучного інтелекту, проте його проміжна, зручна для повсякденного використання, стадія буде створена вже незабаром. А якою вона буде? Які сфери використання буде мати штучний інтелект, здатний вільно спілкуватися з людиною, відігравати деяку роль, вільно змінювати її, виконувати, згідно з роллю, задану складну програму дій? Особливо, якщо в одній системі буде поєднано усі ці технології: по клонуванню голосу, створенню штучних фото та відео людей, веденню зв'язного логічного спілкування. Які налаштування потрібні будуть таким системам, до чого прагнути?

Виклад основного матеріалу

Перш за все, розберемося з очевидним – загальним виглядом майбутньої системи. Зрозуміло, що штучний інтелект, який буде створений із сучасних чатботів, здатний вільно спілкуватися з людиною і обов’язково зможе це робити голосом. Тим більше, що зчитування голосу людини при спілкуванні з наявними чатботами вже використовується широко. Без сумніву, сюди буде включено усі технології виразного, правильного читання, а значить, «розуміння» суті тексту, технології клонування голосу деякої особистості, налаштування цього голосу під себе, технології «розуміння» недосконалої вимови співрозмовника тощо. Вже зараз є сервіси, що дозволяють прочитати текст одним із заданих голосів, вибравши при цьому якесь з доступних емоційних забарвлень цього голосу. Очевидно, що «на борту» майбутньої системи будуть усі налаштування цих технологій, і навіть більше: вибір конкретного голосу якоїсь відомої особистості, створення голосу з деяких параметрів (грубість, тон, діапазон звучання тощо), використання свого голосу, надання голосу емоційної забарвленості зі списку, її автоматичний вибір, залежно від перебігу спілкування, додавання якихось особливостей вимови (від заїкання та шепелявості до конкретного акценту, характерного говору, детального налаштування неправильної вимови, постановки наголосів окремих слів тощо). Голосовий помічник, якого можна детально налаштувати під себе, завжди виграє конкурентну боротьбу у помічника, позбавленого цієї можливості.

Якщо у штучного інтелекту є свій голос, то повинен бути і зовнішній вигляд. Люди прагнуть бачити бодай просте фото співрозмовника, а якщо це інтелектуальний, штучно створений помічник, налаштування його образу буде окремою розкішшю. Знову ж таки, нічого неможливого, технології створення облич людей за заданими параметрами вже існують, існує доволі розвинута індустрія створення фотореалістичної комп’ютерної графіки, технології «оживлення» зображень. Майбутній співрозмовник, помічник, чи яку ще роль буде відігравати штучний інтелект, у будь-якому випадку буде мати певні риси, які можна буде налаштувати. Створюємо редактор, подібний до тих, що використовуються у комп’ютерних іграх, додаємо туди клонування готових образів, їх малювання і на виході маємо анімований образ самого себе, «залізного робота», улюбленого кіногероя, відтвореного з фотографічною точністю відомого політика чи, взагалі, статую Гая Юлія Цезаря із саду у Версальському палаці.

Якщо ми чуємо і бачимо нашого штучного співрозмовника, якщо він здатний підтримувати розмову, якщо його відповіді є адекватними, логічно наділити його певним характером, певними особливостями, що дозволять зімітувати особистість, задати йому якісь особливості поведінки, мету, асоціативні маркери. Тут шлях до ідеалу знову ж можуть підказати сучасні комп’ютерні ігри, де доволі часто є діалоги між персонажами. Їх пишуть люди, найчастіше професійні письменники, що визначають заздалегідь можливі варіанти завершення розмови, варіанти відповідей, доступні для гравця, і, таким чином, формують ігрові діалоги. Уявімо, що ми

маємо систему, яка вміє на непоганому рівні відповідати людині. Логічно, що її обов'язково спробують використати у комп'ютерних іграх. Ця система почне говорити за буквально кожного ігрового персонажа, вона буде вимагати «прокачки» вміння спілкуватися не якоїсь віртуальної ляльки, а безпосередньо людини, яка грає в гру. А для цього потрібні налаштування характеру персонажа, тобто нашого штучного співрозмовника, його цілей, бажань, інформації, що він має, його зв'язків, страхів, механізмів дії на нього тощо. Потрібна анкета штучного інтелекту, текстова, а краще графічна, карта дій, зв'язків, прагнень, локацій, всього, що є важливим. Наприклад, маємо гру про середні віки на кшталт реалістичної Kingdom Come: Deliverance. Перед гравцем стоїть торговець, у якого гравець хоче щось купити. І, як кажуть, тут торг доречний! Якщо в штучного інтелекту є ціль продати гравцю товар за якусь ціну А, й не менше, ніж ціна Б, і штучний інтелект вміє розуміти складні докази, то міні-гра у торгівлю перестає бути нудною. З'являються нові, непередбачувані варіації розмови, величезний пул ситуацій та відповідей, породжених людською фантазією. Так, якщо гравець намагається щось вкрасти, а потім відмовляється від цього бажання, поводить себе підозріло, намагається зайти туди, куди не слід, робить щось дивне, або, навіть, просто перебуває у лавці після часу її закриття, кількість варіацій розвитку подій стає величезною. Від чемного нагадування про час (з останнього прикладу), до, все більш наростаючого роздратування, та виставляння врешті-решт «неадекватної», з точки зору штучного інтелекту, людини за двері. А може ставлення навпаки залишиться дуже терплячим, адже людина не просто сидить у лавці, а ще й поводить себе гідно, виглядає грізно, та й взагалі пообіцяла купити один із найдорожчих мечів, і вже котру годину поспіль вагається, розповідає щось, розпитує і розпитує про товар. Ну як тут відмовити? Відпочинок почекає. Сама гра при цьому отримує просто небачений досі рівень занурення. А інший ігровий персонаж, наприклад, має якісь окреслені заздалегідь страхи, якісь приховані скарби, цілі свого ігрового існування, а отже, його можна залякати засобами звичайної людської мови, у нього можна взяти якесь завдання, що буде відповідати окресленим цілям персонажу, його зиску. Ми отримаємо потенційно нескінченне поле для пригод і творчості. Навіть у недосконалому вигляді подібне матиме величезний успіх і багатомільйонну армію прихильників, що будуть з пристрастю вимагати дослідників розвивати дані технології якнайшвидше. Не важливо, що перед нами: віртуальний помічник у телефоні, система для навчання іноземній мові чи віртуальний диспетчер банку, імітація характеру під себе чи під якісь конкретні цілі, імітація тих чи інших, пов'язаних із зовнішніми факторами, людських особливостей буде невід'ємною складовою проміжної стадії створення штучного інтелекту. Звичайно, механізм побудови «характеру» штучного інтелекту та описаних особливостей зараз виглядає надскладним не тільки з точки зору, як змусити машину все це імітувати, але й як це можна задати, яке середовище для розробки цього потрібно створити, які параметри повинні бути описані, як вони повинні визначати відповіді штучного інтелекту. Але ж початки цього вже є у наявних чат-ботах. Їх розвиток, мабуть, і є найбільшою складністю, що відділяє людство від створення проміжної стадії штучного інтелекту. Саме тому цей напрямок досліджень є найбільш цікавим з точки зору подальших перспектив.

Отже, можна виділити головне. Проміжна стадія штучного інтелекту, подібного до людського, скоріше за все розпочнеться з сучасних чат-ботів та буде синтезом наявних технологій. Вона буде:

- Уміти вільно спілкуватися з людьми, відповідати на їх запитання, формувати власні тощо;
- мати ряд зовнішніх налаштувань, що

складуть її індивідуальність, а саме все, що відноситься до голосу й зображення;

- мати імітацію власного «характеру», складну анкету персонажа та можливість налаштовувати особливості, що будуть накладатися на нейтральну систему спілкування;
- мати гнучку (вірогідно, у графічному вигляді) систему задач, зв'язків, страхів, послідовностей дій, будь-яких логічних маркерів, що потрібні для реалістичного імітування;
- мати можливість вільно й швидко змінювати усі перераховані параметри для конкретного штучного інтелекту, змінюючи тим самим його особистість.

А це, в свою чергу, дозволяє говорити про розробку, в перспективі, деякого стандарту побудови, подібних до людського, штучних інтелектів, розбиття їх структури на окремі, побудовані за однією логікою, пов'язані між собою блоки та вдосконалення далі цих блоків окремо.

У результаті може бути отримана система, що зможе прийти у, як мінімум, наступні сфери господарства:

- Комп'ютерні ігри (не потрібно писати готові діалоги, гравець під час гри тренує свої власні навички спілкування, гра перетворюється на, по суті, тренажер, воістину непередбачувану пригоду).
- Віртуальний помічник, він же – співрозмовник і друг – зручна, гнучка, детально настроювана програма для різноманітних девайсів. (Основа. Сервіс, з яким завжди можна поговорити, дати команди щодо керування гаджетом, у якого можна запитати будь-яку потрібну інформацію, що міститься в мережі інтернет, отримати якусь покрокову інструкцію тощо. Дуже перспективно виглядає помічник, який з екрану мобільного телефона голосом улюбленої зірки відео для дорослих зможе пояснити якусь конкретну річ, наприклад, спосіб користування новим механізмом самообслуговуванням на касі супермаркету та зробить це з урахування налаштованого характеру. Мовляв, бачиш цю велику привабливу прорізь? Міцно візьми купюри в руку й рішуче просунь їх туди одну за одною! Рівно 154 гривні! О так, мій хлопчику! Молодець! :) А в когось у помічниках буде філософ або жартівник, та хто завгодно, будь-яка особистість).
- Автовідповідач, диспетчер для різноманітних установ (функція, що використовується зараз, стане ще більш досконалою).
- Продавець.
- Синхронний перекладач.
- Автоматичний гід, екскурсовод.
- Система спілкування мобільного робота.
- Актор чи просто частина технології з дублювання фільмів (клонування оригінального голосу актора, створення нової міміки, що відповідає зачитаному тексту тощо).
- Підготовка спеціалістів (режим віртуального полону для солдатів, віртуальні допити для слідчих, тренування психологів, психіатрів, перевірка громадян на схильність до розголошення таємниць у розмові, тренування спілкування зі складними підлеглими й начальниками тощо).
- Навчання будь-чому (наприклад, іноземним мовам, взагалі, створення власного віртуального викладача).
- Віртуальна няня, будь-який механізм активного спостереження за чимось (розумна сигналізація на машині зможе не тільки запобігти негарним речам, а й виразно описати, чому не слід робити дурниць, віртуальна доглядальниця за людьми похилого віку зможе їх розважити, вчасно допомогти, або віртуальний таксист, якщо вони залишаться, взагалі буде мало чим відрізнятись від реального).
- Власний тренер і піддослідний кролик (чоловіку потрібно навчитися грамотно спілкуватися з жінкою і навпаки? Віртуальний «винос мозку» до ваших послуг! Тренування здатності вгадувати бажання жінок, навчання «мові натя-

ків», все, що завгодно для самовдосконалення. Особливо цікаво це буде виглядати, якщо можна буде скачувати для тренувань якісь типові налаштування особистостей штучного інтелекту).

- Віртуальний диктор, читець книжок, будь-яких текстів.
- Віртуальний лікар, що зможе провести попереднє опитування пацієнта перед зверненням до лікаря-людини та перебрати на себе частину його функцій, взагалі, пасивно відслідковувати стан хворого.
- Помічник для людей з обмеженими можливостями, синдромом Дауна, із затримками в розвитку (будь-яка допомога, що пов'язана зі спілкуванням).
- Специфічні сфери використання (наприклад, категорія людей, які трохи мають схиблених домашніх улюбленців) могли б поєднати останніх зі штучним інтелектом. Забезпечити їх пристроєм з камерами, різноманітними датчиками для контролю стану тощо.

Висновки

Проаналізувавши розвиток сучасних технологій та успіхи у розробці штучного інтелекту, подібного до інтелекту людини, було сформовано приблизне бачення проміжного стану цього інтелекту на шляху до набуття повноцінних можливостей. Стану, на якому даний штучний інтелект буде спроможний уже значно змінити нинішній світ, у якому його використання стане доцільним і масовим. Виходячи з цього, були сформовані можливі сфери застосування даного проміжного штучного інтелекту.

Література

1. Коломийцев С.Ю. Тест Тьюринга и искусственное мышление в начале XXI века // Человек. 2015. - №4. С 59-68.
2. Shivang V., Lakshay S., Moolchand S. Comparative analysis of chatbots // Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020.
3. Siarohin A., Lathuiliere S., Tulyakov S., Ricci E., Sebe N. First Order Motion Model for Image Animation // 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.
4. Jia Y., Zhang Y., Weiss R.J., Wang Q., Shen J., Ren F., Chen Z., Nguyen P., Pang R., Lopez Moreno I., Wu Y. Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis // 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montryal, Canada.
5. Zhang Xu., Barron J.T., Tsai Y., Pandey R., Zhang Xi., Ng R., Jacobs D.E. Portrait Shadow Manipulation // ACM Trans. Graph., Vol. 39, No. 4, Article 1. Publication date: July 2020.
6. Yu-Lun Liu, Wei-Sheng Lai, Ming-Hsuan Yang, Yung-Yu Chuang, Jia-Bin Huang Learning to See Through Obstructions // arXiv:2004.01180
7. Shu-Yu Chen, Wanchao Su,Lin Gao,Shihong Xia, Hongbo Fu DeepFaceDrawing: Deep Generation of Face Images from Sketches // Siggraph 2020.
8. Karras T., Laine S., Aila T., A Style-Based Generator Architecture for Generative Adversarial Networks // arXiv: 1812.04948
9. Pisarenko V.G., Simulation of the problem of interaction of neurons taking into account the lagging of their interaction .// Cybernetics and Systems Analysis.- Vol. 54, No. 3 (2018). - Published by Springer Science+Business Media, LLC. – P.513- 516.
10. Pisarenko V.G. A New Living Neuronet Model Taking into Account Retarded Interactions Between Neurons. Cybernetics and Systems Analysis 52(6) • November 2016 Published by Springer Science+ Business Media, LLC (USA). https://link.springer.com/article/10.1007/s10559-016-9900-8

References

1. Kolomiytsev S.Yu. Turing test and artificial thinking at the beginning of the XXI century // Man. 2015. - No. 4. C 59-68.
2. Shivang V., Lakshay S., Moolchand S. Comparative analysis of chatbots // Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020.
3. Siarohin A., Lathuiliere S., Tulyakov S., Ricci E., Sebe N. First Order Motion Model for Image Animation // 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.
4. Jia Y., Zhang Y., Weiss R.J., Wang Q., Shen J., Ren F., Chen Z., Nguyen P., Pang R., Lopez Moreno I., Wu Y. Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis // 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montryal, Canada.
5. Zhang Xu., Barron J.T., Tsai Y., Pandey R., Zhang Xi., Ng R., Jacobs D.E. Portrait Shadow

Manipulation // ACM Trans. Graph., Vol. 39, No. 4, Article 1. Publication date: July 2020.
6. Yu-Lun Liu, Wei-Sheng Lai, Ming-Hsuan Yang, Yung-Yu Chuang, Jia-Bin Huang Learning to See Through Obstructions // arXiv:2004.01180
7. Shu-Yu Chen, Wanchao Su,Lin Gao,Shihong Xia, Hongbo Fu DeepFaceDrawing: Deep Generation of Face Images from Sketches // Siggraph 2020.
8. Karras T., Laine S., Aila T., A Style-Based Generator Architecture for Generative Adversarial Networks // arXiv: 1812.04948
9. Pisarenko V.G., Simulation of the problem of interaction of neurons taking into account the lagging of their interaction .// Cybernetics and Systems Analysis.- Vol. 54, No. 3 (2018). - Published by Springer Science+Business Media, LLC. – P.513- 516.
10. Pisarenko V.G. A New Living Neuronet Model Taking into Account Retarded Interactions Between Neurons. Cybernetics and Systems Analysis 52(6) • November 2016 Published by Springer Science+ Business Media, LLC (USA). https://link.springer.com/article/10.1007/s10559-016-9900-8

RESUME

O.S. Koval

An intermediate stage of artificial intelligence similar to human

On the way to perfection, human-like artificial intelligence is likely to go through an intermediate stage, convenient for human use. This stage logically follows from the advances in the development of chatbots and neural networks that are emerging now. It will be characterized by the ability of artificial intelligence to communicate freely with humans, solve logical problems, perform various commands, play a role, change it as needed. At this stage, most likely there will be no question of the real identity of the machine, at the same time its imitation will be available to a wide range. A system of artificial intelligence settings will be created and brought to a convenient look, namely the features of their "character", their tasks, connections, fears, any logical markers needed for a realistic imitation of personality. Full-fledged editors of the voice of artificial intelligence, its photorealistic image, features of these parameters will be created. It is probable to achieve a state when all these things are set according to a single, well-structured scheme, where the same artificial intelligence is trained to communicate, on the fly changes its graphic and voice parameters, personality parameters. Creating such a system can already seriously change the current state of affairs. After all, humanity will receive an extremely convenient, cheap and therefore in demand system that is constantly improving, an assistant that will come to many areas of human life and probably displace people from there. Various virtual announcers, salespeople, simultaneous interpreters, nannies, social workers, answering machines, guides, virtual assistants in all possible areas where communication is needed, virtual doctors who are always ready to help, various trainers, part of simulators, part of mobile robots, artificial intelligence of computer games which are gradually turning into a "matrix" from the film of the same name, all this is our next day. In the form of a program on the devices currently used, a separate static device or a mobile robot, artificial intelligence will definitely come into our lives and who knows how we will meet it.

Надійшла до редакції 20.01.2020